Open Science Grid

Richard:

- What is OSG,
- What has it achieved (both functional capabilities and user base),
- What effort has been involved,
- Brief look at the future

.David: Current and futures would be good.
For the latter, more vision than near term plans would be useful as this is a ten year planning workshop.

Michael/Lothar: We should put emphasis on the collaborative aspect when serving the diverse science and engineering community.

Miron: is in Europe; will comment next week.

Ruth: OSG is diverse in view – these slides and mistakes are mine – others should comment as I go along.

OSG:

Consortium

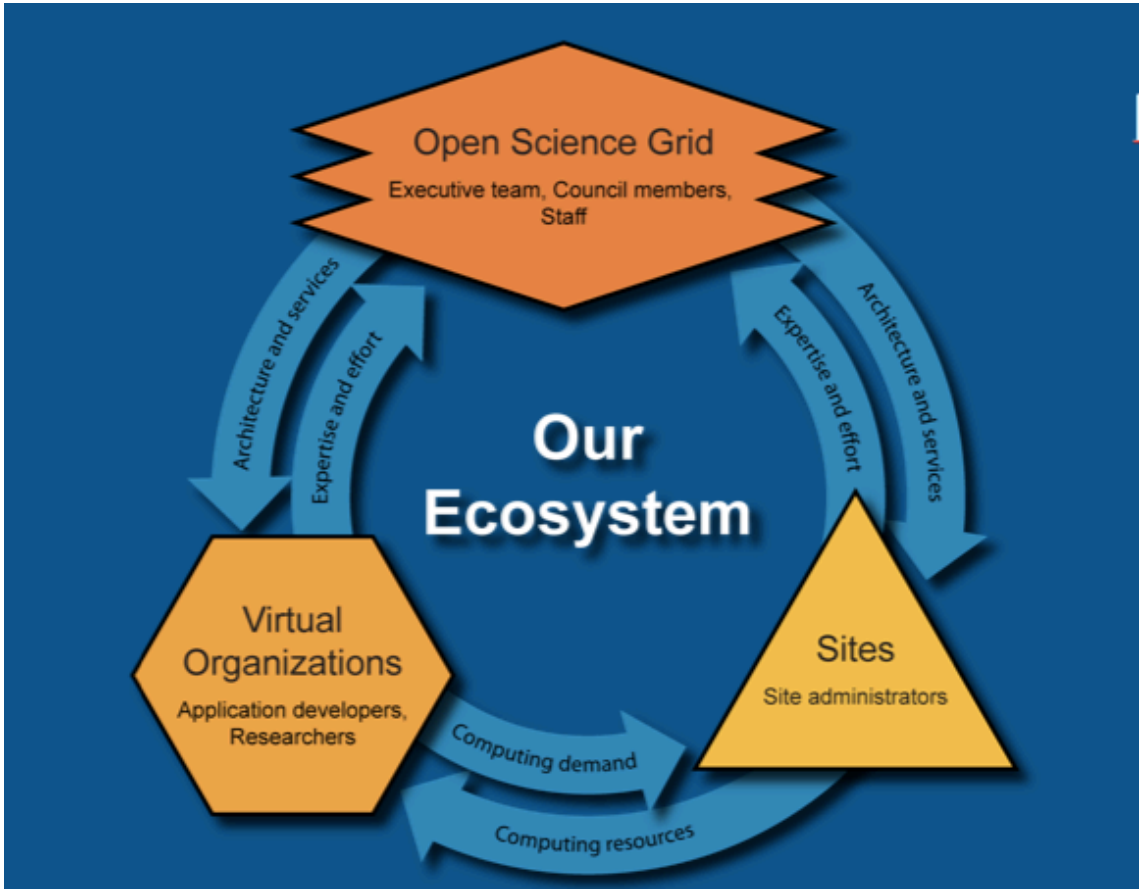
Infrastructures

Project

Satellites

Services:

- Consulting
- Production
- Software



OSG:

Provides and Advances Art of Distributed High Throughput Computing Solutions.

Owns no Resources.

Does not develop Software.

Brings Technologies, Services, Applications, Systems to Production.

People Cooperating:

Scientists – domain and CS

IT administrators

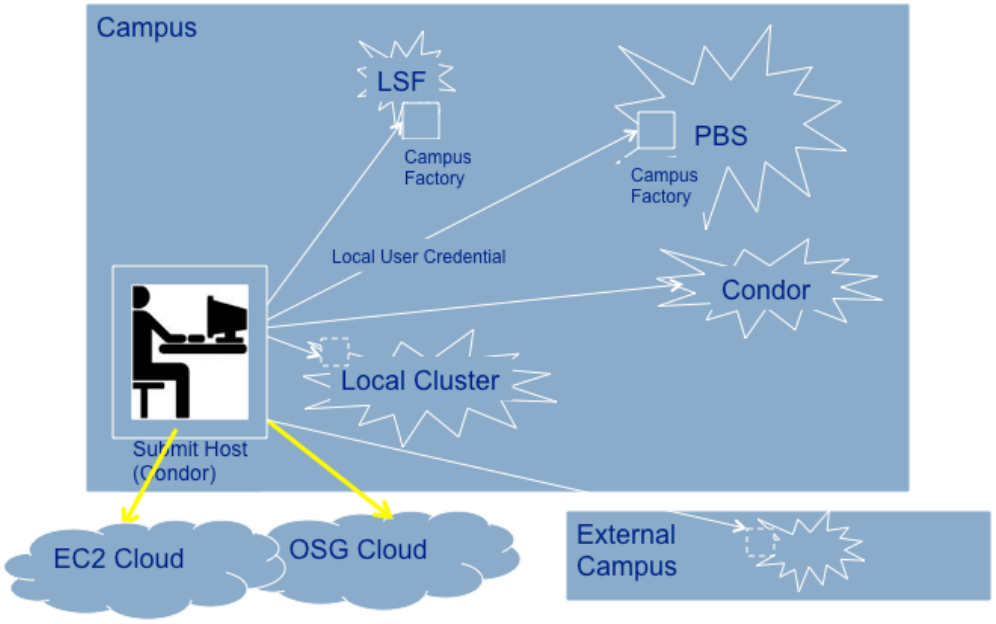
Software developers and support

Security teams

Infrastructures

Production Grid: Set of autonomous, cooperating, shared, CPUs and storage across more than 90 sites in the US and off-shore.

Campus Grids: Set of autonomous, locally shared campus infrastructures, which can also be locally configured to “upload” to the production grid.



Integration Testbed: ~5 shared sites for testing new software and services before release.

Infrastructures

Production Grid: Set of autonomous, cooperating, shared, CPUs and storage across more than 90 sites in the US and off-shore.

Campus Grids: Set of autonomous, locally shared campus infrastructures, which can also be loosely integrated to "upload" to the production grid.

Create Common Grids.

Build Communities.

Retain Individuality.

Evaluate and harden Solutions.

Customer selection of software & services.

Best Practices and principles.

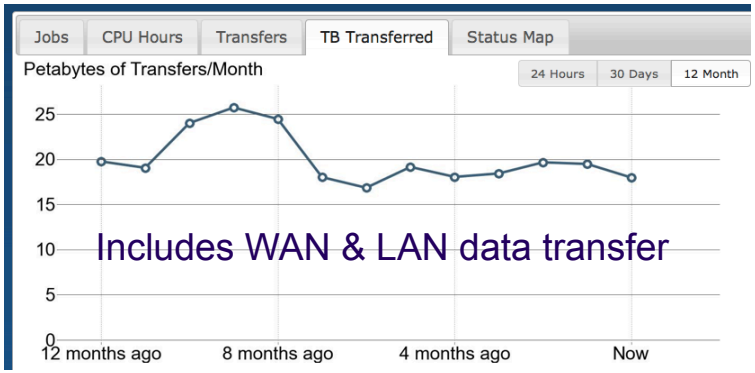
Teach by example – the mature help the youth.

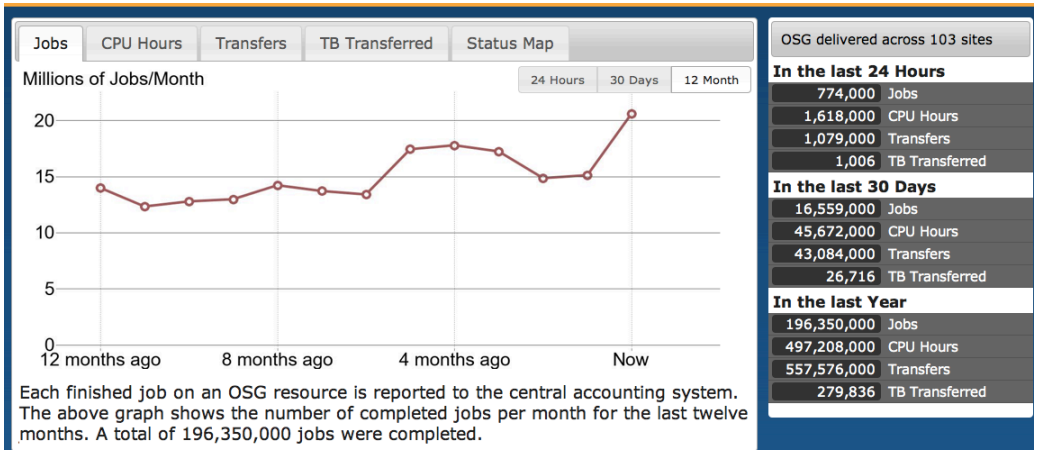
Integration Testbed: ~5 shared sites for testing new software and services before release.

Accomplishments

Delivered MOU'd commitments++ on behalf of US LHC.

Effective Federations with Europe and South America.

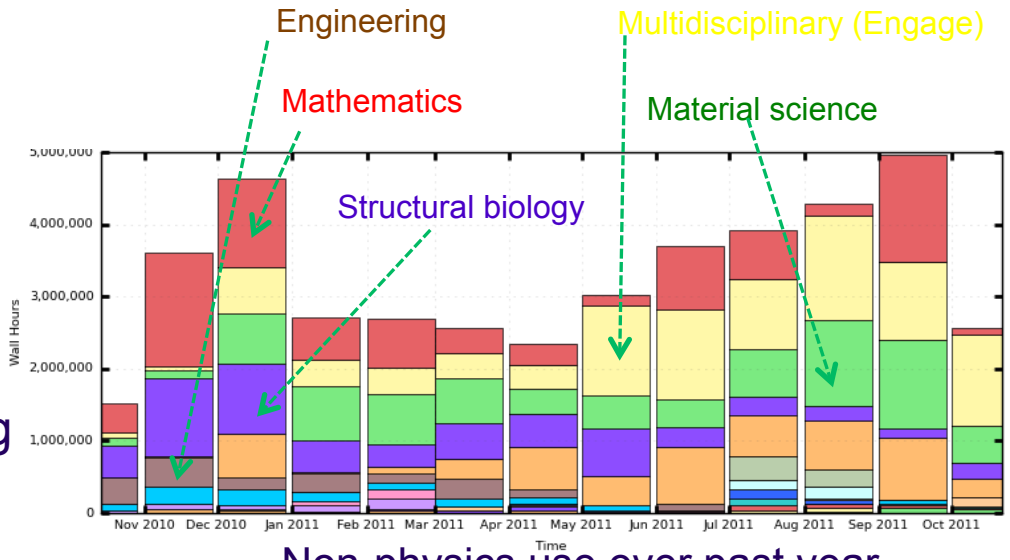




8% CPU cycles used for Non-Physics.

25% CPU cycles used for Non-LHC Physics – Tevatron, LIGO,

Non-LHC usage (25% + 8%) on resources user does not own – making more effective use of available capacity.



Non-physics use over past year

Accomplishments

Increasingly high throughput performance and ease of use across 10s of independent sites based on "job overlays" – pilots, glideins, common workload management system, monitoring and configuration.

High throughput provisioning of massive amounts of data files through standard interfaces to many sites.

Effective sharing of CPU resources increasing return on investments.

Credibility and partnerships with peers in Europe and a foundation for increasing transparency with HPC peers in the US. (OSG invited by XSEDE to be a Service Provider in the NSF XD world.)

An enthusiastic and energetic team spirit including mature communities help the emerging communities.

Non-HEP use over past year

Efforts

Most funding currently ends Mar-Sept 2012

Consortium: all member organizations contribute "in some way".

US LHC contributes most. Software providers contribute support.

Infrastructures: sites each contribute ¼-2 FTE of service support

integrated across application and administration support depending on size of clusters, caches, local user community etc.

Project: OSG project has ~32 FTEs from 2006-2011

consulting ~7; production ~12; software ~10; leadership & admin ~3;

Satellites: total <~10 FTE

High Throughput Parallel Computing	2FTE		CorralWMS (workload management)	1.5FTE
Advanced Network Initiative	1FTE		ExTENCI (TG/OSG)	1.5FTE
Any Data, Any Time, Any Where	1.5FTE		Summer School	0.25FTE
CI-Team/Engagement	1.5FTE			

Efforts

Most funding currently ends Mar-Sept 2012

Consortium: all member organizations contribute "in some way".

US LHC contributes most. Software providers contribute support.

Infrastructures: sites each contribute ¼-2 FTE of service support

integrated across application and administration support depending on size of clusters, caches, local user community etc.

Project: OSG project has ~62 FTEs from 2006-2011 p

consulting ~7, production ~12, software ~16, leadership & admin ~3;

Satellites: total <~10 FTE

Commitment to Sustainability of Production brings commitments to take on support for software/services that are depended upon by the communities but might be otherwise orphaned:

e.g. Bestman from LBNL, Transition of DOE Grids CA, supplements to Condor, Globus/CDIGS-2

High Throughput Parallel Computing	2FTE		CorralWMS (workload management)	1.5FTE
Advanced Network Initiative	1FTE		ExTENCI (TG/OSG)	1.5FTE
Any Data, Any Time, Any Where	1.5FTE		Summer School	0.25FTE
CI-Team/Engagement	1.5FTE			

Concern about and commitment to not abandon Customers once they Depend on the Services

The environment and processes must be there to sustain the software and services for the current set of services.

Evolution and out-sourcing - commercial, agency, institutional - needs, activities, life-cycle costs,

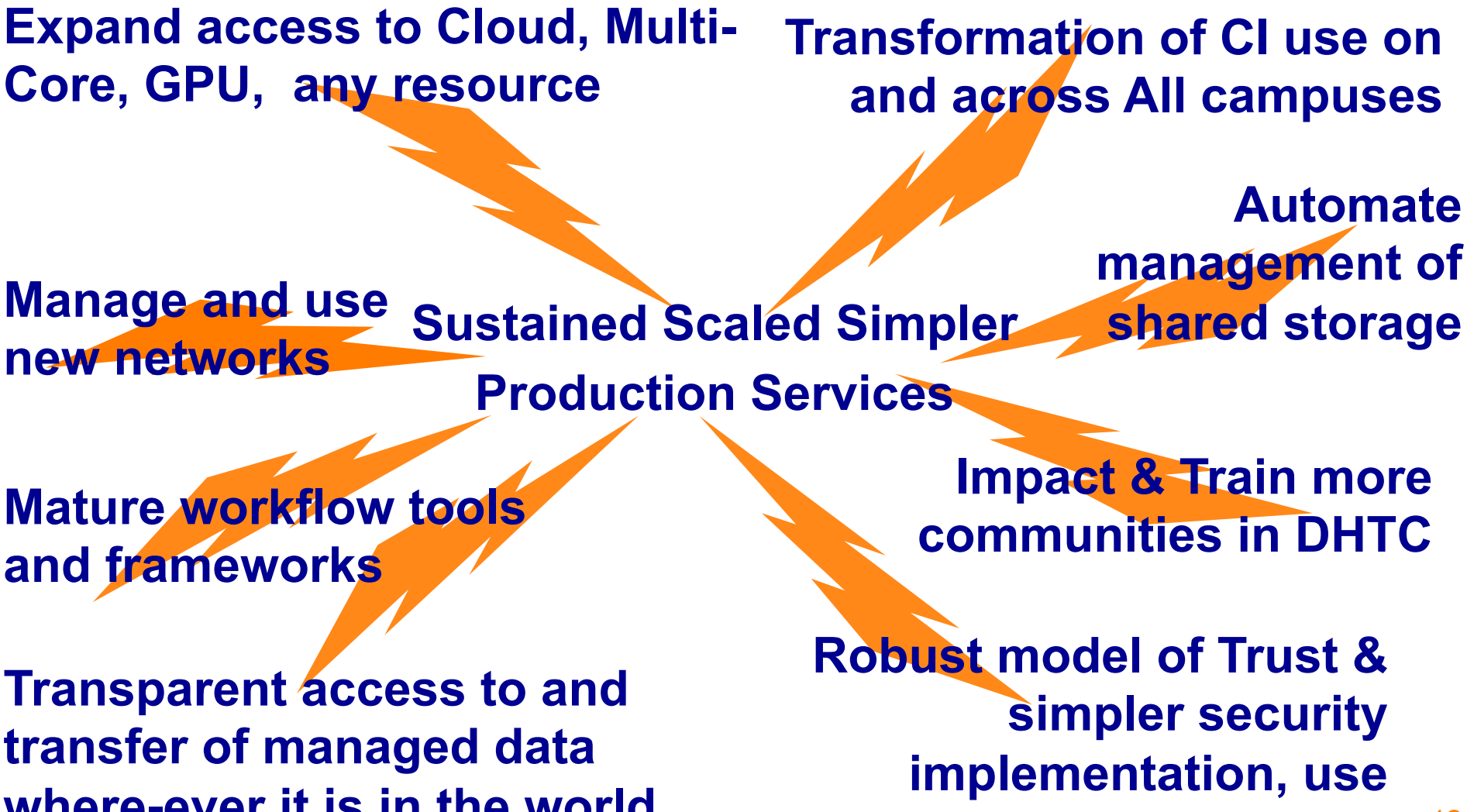
Interoperation and Federation - not only for the large science communities, but for the growing inter-disciplinary and small ad-hoc communities.

Generational Workforce - catering to the new entrants into the workforce and those now entering high-school

Maintain the Open in Open Science - balancing the protection and access needs with the risks and vulnerability practices.

footnote - we are not married to a name

Goals
Satellites to Enhance and Sustain Production



Enhance and Sustain Production Capabilities

With, as, to be ready when, the Communities need them.

Communicating learning & advancing state-of-the art of distributed computing.

Maintaining multi-science multi-agency and multi-science focus.